SIMULATION STANDARD FOR BUSINESS PROCESS MANAGEMENT

John Januszczak

Geoff Hook

Meta Software Corporation
15 New England Executive Park
Burlington, MA 01803, USA

Lanner Group Limited
The Oaks, Clews Road
Redditch, Worcestershire B98 7ST, UK

ABSTRACT

Simulation is considered a key component for business process management suites. Within business process management, simulation can be readily used for both process design, and ongoing improvement. Despite the predictive capabilities of simulation, the lack of wide scale adoption within business process management compared with what might be expected suggests that more can be done to better integrate, and use, simulation with business process management suites. While mature standards exist for the definitions of the business processes, there is a lack of standards for defining business process simulation parameters. This paper describes the current challenges when using simulation for business process management, and shows how a standard for defining business process simulation scenarios would help organizations implementing business process management suites leverage the prescriptive power of simulation. The components of such a standard and how this standard might be extended into a complete process analytics framework will be presented.

1   INTRODUCTION

The term business process management, or simply "BPM" has been used by software vendors, consultants and business people to mean many different, if somewhat related, things. While often overused, it is generally accepted that BPM is a methodology to improve business processes (Global360 2010). A business process management suite (BPMS), is a set of software tools that support the BPM methodology that typically allow organizations who implement them to model, automate, monitor and analyze business processes.

It is helpful to consider business process management as a life cycle of the following activities:

- Design: In this phase, existing and proposed processes are identified and reviewed. Process flows are mapped, the resources required to perform the work are identified, and procedures are considered with an aim to create an efficient set of processes and reduce potential problems.
- Modeling: How the proposed design operates under various scenarios is investigated. Scenarios describe the circumstances in terms of workloads, available resources, and other variables that might be expected in the business. Scenarios considered might include the historical parameters under which processes have been performed, but also "what-if" scenarios that consider potential circumstances for which a business feels it must be prepared.
- Automation: Business process management suites include tools that allow services and line of business applications to be connected to automate the process. These suites may include business rule engines to automatically handle decision points in a process. They typically also provide mechanisms to support the intervention or input of human resources where required by the process.
- Monitoring: Real, near real time, or ad-hoc monitoring of activities and key performance indicators within a business process enables management to react to service interruptions, bottlenecks and other inefficiencies, as well as potential risks and other threats as they occur. In the scope of

- process mining, it also allows discrepancies between the actual running process and the process models to be identified.
- Optimization: Statistics generated by the modeling phase and process monitoring can be analyzed to identify inefficiencies and potential improvements which can be applied to the process design.

Clearly, simulation can be applied directly in the modeling and optimization phases of this life cycle. Computer simulation has been used to analyze business processes since the seventies (Shannon 1975) and simulation is considered a standard analysis technique for operations management. Given the historical use of simulation in business and the BPM methodology described above, it is not surprising that process simulation using real-time, historical and estimated operating parameters is considered a key component of business process management suites (Hill 2007).

2   CHALLENGES FOR SIMULATION

Simulation has a natural and compelling fit in the business process management domain. While this is easy to understand conceptually, it has been argued that many organizations are not using business process simulation in a structured and effective manner (Aalst 2010). The authors believe that this can be attributed to the following factors:

- The inadequate training or skill sets of people tasked with performing simulation analysis.
- The costs of data acquisition.
- The rush to automate.
- Incorrect models.

Looking at each of these issues in greater detail will allow us to understand how a standard for defining business process simulation models will help organizations integrate, and use simulation analysis more effectively. By mitigating these challenges, a standard would allow for a more wide spread adoption of simulation in conjunction with business process management.

2.1   Practitioner Capabilities

The effective use of simulation modeling has traditionally required specialized training. Simulation and other mathematical modeling tools are typically components of an industrial engineering or management science curriculum (Gagne 2009). Simulation analysis is often performed by business analysts when used in the scope of business process management implementations. Without proper training, processes are often modeled incorrectly or not enough data is collected to properly parameterize the model (Aalst 2010).

From a modeling perspective, the authoring of simulation models, the specification of simulation parameters, the execution of the model and the analysis of the results are typically performed in the same software application. In this configuration, the running of simulation scenarios and analysis of the results is usually restricted to those with simulation expertise. When simulation models are specified in a standard way, the running of scenarios and the analysis of the output can be made more accessible to users with less specialized training, even though the initial design of the models might require the skills of a more qualified analyst. As an example: with a standard in place, simulation model parameters could be based on automated queries and aggregations from data sources (for example BPMS event logs) at design time. At run time, simulation parameters could be published into the scenarios via interfaces based on the standard. Simpler user interfaces could also exploit such a standard so that a wider range of users could run ad-hoc what-if scenarios by manually adjusting a subset of parameters in the simulation model (Januszczak 2011).

From a data perspective, a standard for business process simulation scenarios provides a checklist for the types of data required to properly parameterize a simulation model. This provides a guideline for analysts without simulation training to contribute to the acquisition of data required for meaningful models.

At the highest level, a standard allows the design of simulation models, user interfaces for (simulation) analysis, tools that automate the population of simulation model parameters, and simulation engines to be developed independently, by separate solution providers, since they can rely on a common data specification for business process simulation.

2.2   The Costs of Data Acquisition

Computer simulation technology has its roots in manufacturing. While simulation technology has evolved and/or has been adopted for use in business process management, the mindset around its use has not perhaps evolved as quickly. The cost benefit of simulation was traditionally considered when the costs of automation were very high (for example, the automation of the factory floor with expensive, specialized hardware and software systems). When the costs of automation, whether through poor processes or implementation, are perceived to be high, the costs of developing and maintaining meaningful simulation models is often not considered material. However, there are definite costs associated with simulation which become increasingly evident when the costs of automation are much lower. Besides the need for specialized talent described in the previous section, simulation models require a typically large data set of parameters or inputs such as volumes and their arrival patterns, task durations, resource availability and so forth to ensure the model generates meaningful and accurate results. Furthermore, if the simulation model is to ever be used again, it is likely that these parameters will have to be actively updated on a regular basis.

A standard for defining business process simulation scenarios allows other process aware systems to exchange data with simulation software that supports the standard. Examples of process aware information systems include BPMS, enterprise resource planning (ERP) systems, process centric middleware, some line of business applications, and call center systems (Aalst 2009). This encourages both systems and simulation vendors to support such a standard, and the cost of acquiring the data necessary for robust simulation models is driven down by the availability of data in a readily consumable format. In the current state of affairs, available artifacts are often not being used to update simulation model parameters in an effective manner: event logs and other system data may be used to estimate simulation parameters, however, the parameters are often represented by some kind of distribution, instead of being used directly (Januszczak 2011). This estimation is often performed at model design time and not subsequently updated by using more current data (Aalst 2009). A standard allows for the direct use of this data in the simulation model reflecting the most up to date state of the system.

2.3   The Rush to Automate

Many BPM vendors and consultants consider it a mistake to spend too much time on process analysis at the expense of delaying automation. When the major component of a BPMS is automation and the monitoring of data generated by such automation, the specific real return on investment for the software solution will only be realized once a process has been automated. There is also a justified fear of "analysis paralysis" where the attempt to gold plate a process design or model has resulted in nothing more than binders full of flowcharts sitting on shelves collecting dust. When process design and modeling cannot keep up with real world processes that change frequently, process models quickly become out of date.

Even if there is a rush to automate, it must be remembered that simulation is not purely a design tool. Continuous process improvement and predictive capabilities are possible when simulation models are consumers of automation data. This is perhaps one of the strongest cases for a business process simulation standard: when systems are already automated, data acquisition and simulation model maintenance costs come way down if there is a standard format in which to share the data generated by automation. In some cases it might make more sense to simply employ simulation after automation, not before it. As shown in Figure 1, the costs of simulation are driven down while the accuracy, and therefore predictive ability, of simulation models increase. Why is this so? Automation collects the data required by simulation models. Automating the collection of data and in some cases the process definition itself through process discov-

ery reduces the need for specialized talent. Automation also provides a means of providing the data required to maintain the models for ongoing use (both simulation parameters and process model extensions). Lastly, the data collected from an automated system will most likely be more accurate than data compiled by hand for static, steady state simulation models, which leads to better simulation results.
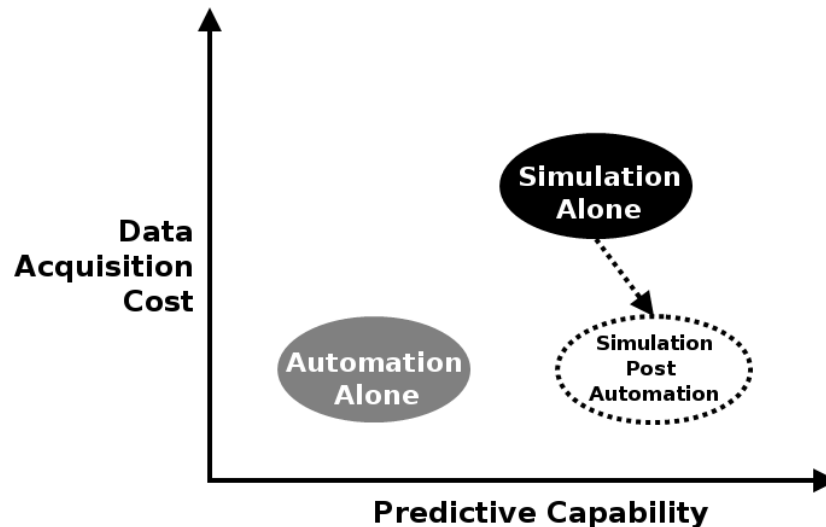

Figure 1: Data acquisition costs are low when automation has been implemented, however, automation alone has little or no predictive capability when compared to simulation. When simulation is a consumer of the data generated by automation, the data acquisition costs are lowered and the accuracy of simulation results improve.

2.4 Incorrect Models

At times, simulation models can contain more detail than is necessary for the issue being investigated. The converse is also true: simulation models can also be oversimplified. Two types of simplifications often negatively impact the use of simulation in business process management: simple resource models and empty initial states. Simulation models often do not reflect that fact that people have multiple skills, work in multiple processes, that they have different proficiencies depending on the activity or type of work being performed, or that resources are not continuously available. Furthermore, many simulation models assume an the initial state where the model is empty. While this can reflect reality in cases where there is no work carried over from the periods preceding the scenario start time, this is usually not accurate. A common technique is to use a warm up period and disregard the simulation results from the initial periods. Because of this, simulation cannot be used for any kind of transient analysis except for the special case where the actual process under study starts from an empty state.

A standard for defining business process simulation scenarios addresses both of these concerns. A standard would provide for the definition of a rich resource model allowing for the specification of specific resources and their availability and efficiencies across the various roles and activities to which they may by assigned instead of simple models such as resource level by role or function. A standard would also support various resource usage patterns such as direct use, role based use and contingent allocation, amongst other common patterns (Russell 2004). Such a standard allows tools that naturally capture the attributes of a rich resource model (such as scheduling systems, time and attendance systems, and Human Resources Information Systems) to interchange, and update, this data with simulation models to more accurately reflect the resourcing configuration of the business processes being simulated. Similarly, the current state of production and/or work flow systems could be used to initialize the simulation model by specifying the process state at the simulation start time using the standard.

3   PROPOSED STANDARD FOR BUSINESS PROCESS SIMULATION SCENARIOS

Simulation models used in business process management are characterized by a similar set of parameters. Understanding the common elements of all business process simulation model parameters allows us to identify what information is required to adequately describe a business process simulation scenario and therefore to enumerate the components of a standard definition and interchange format. The components can be broken down as follows:

- Scenario meta data and context
- Process descriptions
- Events
- Resource model
- Activity parameters
- Tool extensions

The components of a standard for defining business process simulation scenarios are illustrated in the meta model shown in Figure 2. The components of the meta model which reflect the common components of business process simulation scenarios are described in the sections that follow.
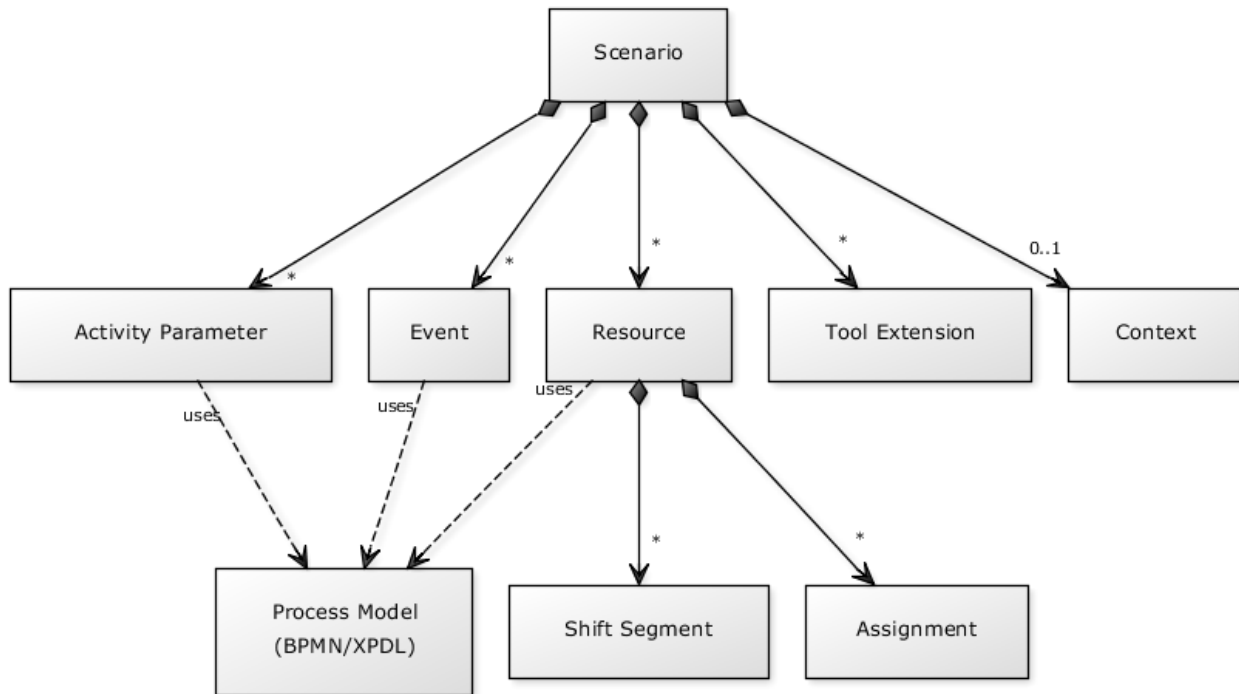

Figure 2: Business process simulation scenario meta model.

3.1   Scenario Meta Data and Context

Scenario meta data describes the scenario itself, such as a unique identifier for the scenario, the author, the tool that generated the scenario definition, as well as scenario creation and modification timestamps.

Table 1: Scenario meta data.

Element	Description
Name	The scenario name or id. This is a unique identifier for the scenario.
Author	The author of the scenario parameters.
Editor	The software tool or service used to define the simulation scenario parameters.
Creation Timestamp	The date and time when the scenario was first created.
Modified Timestamp	The date and time when the scenario was last updated.

The context includes the start time and duration of the simulation and other global settings as described in Table 2. It would also be useful to include a mechanism to reusing a scenario, so that other scenarios or experiments could be defined by referring to an existing scenario definition.

Table 2: Scenario context.

Element	Description
Start Time	The value for this element specifies the start time of the simulation.
Duration	This represents the scenario length defined in terms of the specified clock units. For example, if the clock units are MINUTE, then a duration of 2880 represents a scenario that runs for 2 days.
Clock Units	DAY, HOUR, MINUTE, or SECOND. The default unit of measure for time expressions.
Scheduling Method	CONCURRENT or EXCLUSIVE. Concurrent scheduling represents a global state where all resources can play any role to which they are assigned when they are available for productive work. Exclusive scheduling indicates that the roles which a resource can play at any given time are specified explicitly in the resource's schedule.
Resource Selection	LIST, RANDOM, PERFORMANCE. List means that an attempt to acquire a resource is done in the order by which they are specified (for a given role). Alternatively, the resources can be acquired randomly for a given role or in order of performance where the resources that have the highest performance in a given role are acquired first depending on availability.

3.2 Process Descriptions

Often referred to as the control flow (Aalst 2010), the process definitions describe the ordering of activities in a business process. Activity order may be sequential, or depend on splits, joins and/or loops in the flow of work. Since resources may participate in multiple business processes, it is often convenient to include multiple processes in a simulation scenario.

The definition of business processes is already covered by existing standards, specifically Business Process Modeling Notation (BPMN) and the XML Process Definition Language (XPDL) within the business process management domain. Therefore, a standard for defining simulation models would leverage these existing standards by loosely coupling the rest of the simulation parameters to elements in the

process definitions represented in BPMN or XPDL. For example, an activity parameter like activity duration, would refer to activities in a corresponding BPMN diagram or XPDL document.

3.3 Events

Within a business process scenario, events occur that affect the flow of the process. An event can initiate a process, occur during a process flow, or terminate a process flow. From a simulation modeling perspective, an event represents the instantiation, injection (or removal) of a process instance or "token" in the simulated work flow. Many simulation tools would refer to these as the arrival patterns (and exits) from the model (note that the removal of tokens from a simulation model may be implicit depending on the specific tool). In a simulation model, process instances can be instantiated at any point in the process, and therefore could be associated with a process start event or an intermediate event in BPMN or XPDL. To initialize the state of the process at the beginning of a scenario's period, the standard would also allow for the instantiation of tokens at process activities (unlike BPMN). Also unlike BPMN, there is no need to differentiate simulation scenario events by trigger type. Table 3 describes the elements of a scenario event.

Table 3: Scenario events.

Element	Description
Event ID	A unique identifier for the event.
Process Name	The process name or id (e.g. as defined in the BPMN or XPDL process descriptions).
Data Source	Specifies the source system or scenario of the event data.
Destination	The element/node in the process model where the event takes place. Events may occur at places other than the start of a process to represent the work in progress at the start time of the scenario. In general, this is a reference to an event or activity in the process model.
Event Time	The date and time at which the injection into the process takes place.
Occurrences	How many times the event occurs. This could be a statistical distribution.
Frequency	The time between occurrences. This could be a statistical distribution.
Quantity	Allows multiple events with the same attributes to occur at the given event time. This could be a statistical distribution.
Field Values	Data field values for the event. Since events instantiate tokens in the simulation, the field values can be used to describe the attributes of the injected token. To initialize the state of the process at the beginning of the simulation period, these values may include the resource currently working on the token.

3.4 Resource Model

The resource model describes the specific resources, human or otherwise, used to perform the work in a business process. Resources may be assigned to multiple roles (where roles typically represent skill sets or a specific type of expertise required to perform a certain set of activities) to reflect multiple skill sets. Roles may be loosely coupled to the roles or performer types described in BPMN or XPDL process definitions. Resources may be assigned to a role, a specific set of work according to specific characteristics of

the work, to specific activities or some combination of the preceding. Resources are also characterized by their availability to perform work in a given scenario, for example, by a set of shifts or uptime.

Table 4: Scenario Resources

Element	Description
Name	A unique identifier for the resource.
Department	Associates the resource with an organizational unit. This could be used for reporting purposes.
Quantity	A resource could represent multiple instances of a resource with identical attributes.
Cost	Resources often have costs associated with them (e.g. salary, hourly rate, etc.). This field can be used to calculate costs by resource usage or availability.
Assignments	Specifies that type of work the resource is assigned to work on and their proficiency at processing that type of work. This is how a resource is assigned to a specific role, a specific class of work, and perhaps a specific set of activities. An assignment represents a role and/or list of data field values that must be true in order for the resource instance to work on those items. For example, a resource performing role X might only work at location Y and on work associated with customer Z. Proficiency may vary by role, customer, or any field value combination.
Shift Segments	Shift segments are used to describe the availability of resources during the scenario. Furthermore, a shift segment can optionally be dedicated to a particular assignment (see above).

3.5   Activity Parameters

Activity parameters describe how work is processed at a specific activity referred to in a BPMN or XPDL process definition. If necessary, there may be multiple sets of parameters for a given activity to represent situations where activity parameters vary from one kind of work to another when processed at the same activity.

3.6   Tool Extensions

Simulation engines which will simulate the scenario defined using this standard may have useful options which should be persisted. For example, a specific tool may allow a seed value to be set so that two simulation runs are forced to use the same random numbers. Alternatively, other tools may support custom initialization procedures to be performed. Whatever the particular case, the standard should allow for vendor specific extensions to be defined within the standard's schema.

Table 5: Scenario activity parameters

Element	Description
Activity Name	A unique identifier for the activity that references an activity in the process model.
Field Values	Describes the type of work for which the activity parameters apply. For the same activity, different parameters may apply depending on the attributes of the work item being processed at the activity.
Priority	The relative ranking of activities. The performing of one activity could take precedence over the performing of another.
Resource Acquisition	Specifies how, and which, resources are acquired to perform the activity. This could be the specification of a role or set of roles, a specific resource or set of resources, or some kind of logic or expression that dynamically specifies the acquisition of resources according to the work being processed and/or state of the scenario.
Duration	Specifies the duration of performing the activity once any resources specified are acquired. This could be an absolute value, some kind of statistical distribution, or some logic or expression that specifies the duration based on the work being processed, the resources being used or the state of the scenario.
Additional Operations	Used to specify any custom logic or processing that must be executed prior to any acquired resources being released.
Resource Release	Specifies how, and which, resources are released when the activity is completed.
Weights/Conditions/ Routing Information	The routing of output produced by an activity to subsequent activities may depend on relative percentages or weights, or on some other condition or set of conditions.

3.7 More Reading

It should be noted that the "standards" construct tabled in this paper is an early draft of what shape such a standard might take, specifically illustrating the data that such a standard would contain. Describing a complete specification of such a standard for defining business process simulation scenarios is beyond the scope of this paper, however, an early draft of a specification based on the contents of this paper can be found at the Sim4BPM website (Sim4BPM 2010). This draft specification has been used to generate a preliminary xml schema (Sim4BPM 2010) and client code libraries (Sim4BPM 2010) which can be used by applications to interact with the schema. It has also been shown that a schema based on this draft standard can be readily used to develop a RESTful web services API (Januszczak 2011).

4 PROCESS ANALYTICS FRAMEWORK

Simulation models are just one possible consumer of the business process scenarios that the proposed standard sets to define. For example, one can imagine activity based costing models, or workforce management applications using the same set of data. The primary driver behind this paper is business process simulation, however, simulation is a special use case for the business process scenarios that would be described by this standard. Business process simulation output, data that was not discussed in this paper, is also related to the ideas that have been outlined, and yet again simulation is a special case. Other tools, specifically the automation engines within business process management suites, generate a set of results

of the same type as simulation (for example, the comparison of actual and what-if results depends on the side by side viewing of real process results and simulation results). The standard described in this paper could be extended to include other process based tools and their output as shown in Figure 3. This is an area for future research.
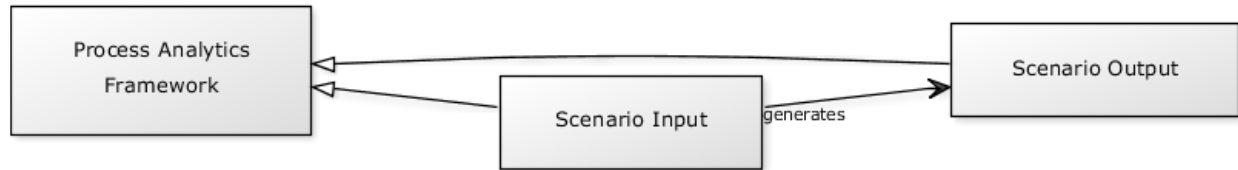

Figure 3: A process analytics framework would encompass scenario parameters (inputs) and scenario outputs.

5   CONCLUSION

Simulation provides powerful predictive capabilities for business process management. However, creating and maintaining simulation models can often be time consuming and expensive. Since the types of information required to define business process simulation scenarios is well known, and consistent across existing business process simulation solutions, the development of a standard schema for the specification of business process simulation scenarios would allow for the exchange of data between business process management suites and other process aware information systems. This would lower the costs of business process simulation and allow organizations implementing business process management to better leverage the predictive and prescriptive features of simulation analysis.

ACKNOWLEDGMENTS

REFERENCES

Aalst, W.M.P. van der. 2010. "Business process simulation revisited." In *Enterprise and organizational modeling and simulation (6h International Workshop, EOMAS 2010,* held at CAiSE 2010, Hammamet, Tunisia, June 7-8, 2010. Selected Papers, edited by J. Barjis. Lecture Notes in Business Information Processing, Vol. 63, 1-14. Berlin: Springer.

Aalst, W. M. P. van der, J. Nakatumba, A. Rozinat, and N. Russell. 2009. "Business Process Simulation: How to get it right?" In *International Handbook on Business Process Management*, edited by J. vom Brocke and M. Rosemann. Berlin: Springer-Verlag.

Gagne, D. 2009. "Modeling and Simulation in Business Process Management". Last modified June. http://www.slideshare.net/dgagne/modeling-and-simulation-in-business-process-management.

Hill, J. B., M. Cantara, E. Deitert, and M. Kerremans. 2007. "Magic Quadrant for Business Process Management Suites, 2007." Gartner RAS Core Research Note G00152906, Gartner, Inc., Stamford, Connecticut.

Januszczak, J. 2011. "Simulation for Business Process Management." In *BPMN 2.0 Handbook*, edited by L. Fischer, 43-57. Lighthouse Point, Florida: Future Strategies.

Global360. 2010. "BPM Primer: People and Processes." Last modified November 5. http://global360.com/xres/uploads/resource-center-documents/110510_g360_wp_BPM_Primer.pdf.

Russell N., A.H.M. ter Hofstede, D. Edmond, and W.M.P. van der Aalst. 2004. "Workflow Resource Patterns." BETA Working Paper Series, WP 127, Eindhoven University of Technology, Eindhoven, Netherlands.

Shannon, R. E. 1975. *Systems simulation: the art and science*. Englewood Cliffs, NJ: Prentice-Hall.

Sim4BPM. 2010. "Client Libraries". Last modified October 6. http://sim4bpm.com/client-libraries/

Sim4BPM. 2010. "Sim4BPM Schema". Last modified October 4. http://sim4bpm.com/schemas/scenario.xsd

Sim4BPM. 2010. "Specification". Last modified October 4. http://sim4bpm.com/specification/.

AUTHOR BIOGRAPHIES

JOHN JANUSZCZAK is a process modeling and simulation expert who designs solutions that optimize process, staffing and resource levels for organizations with complex processes. John has a degree in astrophysics from the University of Toronto and has subsequently been engaged in software development and business process transformation for 20 years. John has worked in a broad spectrum of industries from telecommunications to financial services at such companies as Rogers Communications and Assurances Générales de France. Currently, John is working on the integration of process modeling and simulation technology with workforce management applications, as well as developing process analytics and simulation solutions as RESTful web services. John is currently employed by Meta Software Corporation. His email address is jjanuszczak@metasoft.com.

GEOFF HOOK has worked in the field of Operational Research and Simulation for 30 years after graduating from the University of Warwick Business School and gaining an MSc in the Social Aspects of Science and Technology from the University of Aston in Birmingham. Geoff is one of the founders of the Lanner Group who specialize in simulation software and services, developing products such as WITNESS and L-SIM which is a simulation engine for BPMN integrated within a number of world leading Business Process Modeling solutions. At the Lanner Group he has held a number of senior positions in consulting as well as in new product and solution design. Geoff Hook is a member of the UK Operational Research Society and sits on the societies' Awards Committee including adjudicating on the Tocher award for the best simulation paper in the Journal of Simulation. He is also a member of the Warwick Business Schools' MSc Course Advisory Board. His email address is ghook@lanner.com.